
Bibliographical Data Workflows: discover, analyse, and improve them

Vojtěch Malínek*¹ and Tomasz Umerle*^{†2}

¹Institute of Czech Literature, Czech Academy of Sciences – Czech Republic

²The Institute of Literary Research of The Polish Academy of Sciences – Poland

Abstract

Bibliodata involves us all

Bibliographical data serves to identify documents crucial for the humanities - from books, journals, or articles, to grey literature and web-resources. All researchers perform bibliodata activities: we (co)create, and (re)publish documents, (co)produce bibliographical descriptions, (re)use them, or depend on citation metrics. Also, we all take advantage of the services which depend on bibliodata quality, from library catalogues, through scholarly repositories and digital libraries, to bibliodata managers. Lastly, bibliodata research has heavily influenced DH landscape, with contributions in literature (F. Moretti, M. Jockers, K. Bode), book history (S. Burrows, A. Montoya, M. Tolonen), correspondence network studies (H. Hotson, D. Edelstein, T. Wallnig) or citation studies (S. Peroni).

Bibliodata workflows and scholarly primitives

Collaborative bibliodata workflows are repeatable patterns of co-dependent activities involved in bibliodata curation, and research.

Analysis of research activities that constitute those workflows is crucial to discussions about a contemporary set of scholarly primitives, and the relevance of the notion itself.

During our workshop we want to investigate bibliodata activities performed by the participants on a day-to-day basis, and highlight their role in the research workflows. We will try to assess whether the actual bibliodata activities are well-supported by existing bibliodata research infrastructure, but also discuss the nature of current bibliodata RI's limitations.

Workshop goals

We aim to:

1. present basic information on bibliodata,
2. survey the participants' bibliodata activities, and investigate their role in workflows,

*Speaker

[†]Corresponding author: tomasz.umerle@ibl.waw.pl

3. analyse bibliodata workflows through critical analysis of existing patterns.

Programme:

1) **Introduction**

2) **Bibliodata and scholarly primitives - participants survey:**

- a) what bibliodata they work with and what activities they engage in,
- b) which tools, and services they use - which stakeholders and research approaches they interact with,
- c) how bibliodata activities relate to the participants' work, what value and meaning they assign to it.

3) **Critical investigations of crucial bibliodata workflows:**

Short presentations on topics that will pinpoint crucial challenges to contemporary bibliodata workflows. Topics include:

"Analysing and enriching bibliodata with AVOBMAT (Analysis and Visualization of Bibliographic Metadata and Texts)"

"Author disambiguation using Linked Open Data"

"Cultural phenomena seen through the bibliographical data research"

"How to create the workflow to merge national subject bibliographies?"

"Online resources for bibliodata discovery: usage, limitations and desiderata"

"Online resources for the discovery of bibliodata"

"The role of subject bibliographies in the contemporary research ecosystem"

"Web archiving and bibliographies of internet contents"

"Zotero etc. help you a lot - but where are its limitations?"

Participants will be able to indicate the topics they are interested in through a **pre-workshop survey** which will

4) Conclusion: Based on the critical investigations of identified bibliodata workflows, we would like to think outside of the box: how can we improve the limitations of the current bibliodata workflows? Who should be involved in this process?

Target audience

No previous experience of bibliodata is expected. The individual level of expertise can be stated during the pre-workshop survey. We welcome participants from any field interested in bibliodata.

Keywords: bibliographies, bibliographical data, metadata, data analysis, data quality assessment